Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2012

A. Yourtchenko
cisco
March 5, 2012

Revealing hosts sharing an IP address using ICMP Echo Request
draft-yourtchenko-nat-reveal-ping-00

Abstract

When an IP address is shared among several subscribers -- with a NAT or with an application-level proxy -- it is impossible for the server to differentiate between different clients. Such differentiation is valuable in several scenarios. This memo describes a technique to differentiate TCP and UDP clients sharing an IP address. The proposed method uses an ICMP Echo Request packet, which allows for more information about the user mapping to be transmitted than in the case of using the TCP option - and allows the use with UDP and other protocols.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction 3
2. Terminology 3
3. Description 3
3.1. Operation of Address Sharing Device 4
3.2. Encoding the information into the ICMP Echo Request 4
3.3. Operation of the TCP Server 4
4. Interaction with the transport layer protocols 5
4.1. Upstream NATs and Load Balancers 5
5. Interaction with TCP SYN Cookies 5
6. Security Considerations 5
7. Acknowledgements 6
8. IANA Considerations 6
9. References 6
9.1. Normative References 6
9.2. Informative References 6
Appendix A. Change History 7
Author's Address 7

1. Introduction

The transport-layer proposal mentioned in [I-D.wing-nat-reveal-option] has one drawback: a relatively small amount of information that can be transmitted. This is caused by the fact that the TCP option space is very scarce.

Another potential problem is blocking of the new TCP option by the middleboxes, which, as [I-D.abdo-hostid-tcpopt-implementation] shows a noticeable failure rate of 2.6%

This document describes a mechanism where the address sharing device encapsulates the necessary differentiating information into an ICMP Echo Request packet that it sends in parallel with the initial session creation. The information included in the ICMP Request Data portion describes the five-tuples as seen on both of the sides of the translating device. This allows the server to differentiate different internal addresses. At the same time, since the data travels in ICMP packets, even if they are blocked on the way, the user connection does not have to block.

An analysis of other techniques is available in [I-D.boucadair-intarea-nat-reveal-analysis].

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC2119 [RFC2119].

subscriber: the client accessing an address sharing device, who is responsible for the actions of their device(s). This might be an individual handset (with mobile devices), a home Internet connection, a small-medium business Internet connection, a University dormitory room, an individual employee of a company, or the company itself.

3. Description

This proposal suggests to initiate the ICMP Echo Request / Echo Response exchange with the target for each of the new sessions that are being created. The data portion of the ICMP Echo Request packet will contain the necessary information about the connection - internal and external five-tuples, as well as any other information that the translation device considers useful to share.

The transactional nature of the ICMP Echo/Echo Reply sequence allows

to assert the fact of the remote server (or ICMP proxy thereof) receiving the information - thus, this approach allows reliable transfer of translation information to the target.

3.1. Operation of Address Sharing Device

Upon the creation of the new address mapping, the address sharing device initiates the new ICMP exchange with the target. This exchange happens in parallel with the main connection establishment. In case of not receiving the ICMP Echo Reply, the address sharing device MUST retransmit the echo requests several (exact value TBD) times with exponential backoff.

The source of the ICMP Echo Reply MAY be a separate address on the address-sharing device, dedicated to these ICMP exchanges.

```
     TCP client           address sharing device      TCP server
          |                         |                     |
          |---TCP SYN-------------->|                     |
          |                         |----TCP SYN--------->|
          |                         |--ICMP Echo Request->|
          |                         |<---TCP SYNACK-------|
          |<--TCP SYNACK------------|                     |
          |                         |<-ICMP Echo Response-|
          |---TCP ACK-------------->|                     |
          |                         |--TCP ACK----------->|
          |                         |                     |
```

3.2. Encoding the information into the ICMP Echo Request

A strawman proposal is to use the payload within the Echo Request similar in format to the one described in [I-D.shen-traceroute-ping-ext], with a different magic number. This has the advantage of reusing the code and allowing for some forms of authentication of the NAT devices.

3.3. Operation of the TCP Server

The TCP server identifies the ICMP Echo Request as a special one by inspecting the payload and matching the included outside five-tuple with one of its active connections. After the processing of the Echo Request packet the server sends back the Echo Reply packet with identical contents.

Note that the out-of-band nature of the proposed signaling allowes

multiple scenarios of implementing the server-side handling. An early prototype implementation is in progress using libipq on Linux, which would delay the inbound SYN segments for a configurable interval, in the hope that the ICMP Echo Request with the translation details arrives shortly.

Another implementation could employ some more sophisticated processing - e.g. intercept the SYN segments only from hosts who according to certain heuristics are misbehaving - thus, avoiding any delay for the well-behaving hosts.

4. Interaction with the transport layer protocols

This section discusses the pros and cons of using a separate channel to discriminate the internal 5-tuples.

4.1. Upstream NATs and Load Balancers

The upstream translators may not understand the contents of the packet, and might simply translate it as another ping packet exchange. TBD: the data format needs to provision for detection of this. This item needs further consideration, specifically how to cascade the multiple translators in a chain.

However, the extra address translator north of CGN-style one is rather unlikely.

5. Interaction with TCP SYN Cookies

TCP SYN cookies [RFC4987] are commonly deployed to mitigate TCP SYN attacks, which have some side effects - the ICMP Echo packet containing the 5-tuple mapping information may not match an existing TCP connection on the server. However, in this case the flow of operation of neither TCP SYN Cookies nor ICMP Echo Request is disturbed - the host can simply respond as normal. TBD: one way is for the server to defer sending Echo Reply if there is no matching connection - this will cause the Address Sharing Device to keep retransmitting the Echo Request, and if the connection is legitimate, then eventually the Echo Request will match a newly established connection.

6. Security Considerations

An attacker might use this functionality to appear as if IP address sharing is occurring, in the hopes that a naive server will allow

additional attack traffic. TCP servers and applications SHOULD NOT assume the mere presence of the functionality described in this paper indicates there are other (benign) users sharing the same IP address.

7. Acknowledgements

Thanks to Dan Wing for the discussions, the reviews of early versions of the draft, very helpful suggestions on the text and the nice ASCII art.

8. IANA Considerations

None.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, June 2010.

9.2. Informative References

[I-D.abdo-hostid-tcpopt-implementation]
Abdo, E., Boucadair, M., and J. Queiroz, "HOST_ID TCP Options: Implementation & Preliminary Test Results", draft-abdo-hostid-tcpopt-implementation-02 (work in progress), January 2012.

[I-D.boucadair-intarea-nat-reveal-analysis]
Boucadair, M., Touch, J., Levis, P., and R. Penno, "Analysis of Solution Candidates to Reveal a Host Identifier in Shared Address Deployments", draft-boucadair-intarea-nat-reveal-analysis-04 (work in progress), September 2011.

[I-D.despres-intarea-4rd]
Despres, R., Matsushima, S., Murakami, T., and O. Troan, "IPv4 Residual Deployment across IPv6-Service networks (4rd) ISP-NAT's made optional", draft-despres-intarea-4rd-01 (work in progress), March 2011.

[I-D.ietf-intarea-shared-addressing-issues]
Ford, M., Boucadair, M., Durand, A., Levis, P., and P. Roberts, "Issues with IP Address Sharing", draft-ietf-intarea-shared-addressing-issues-05 (work in progress), March 2011.

[I-D.ietf-mptcp-multiaddressed]
Ford, A., Raiciu, C., Handley, M., and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses", draft-ietf-mptcp-multiaddressed-04 (work in progress), July 2011.

[I-D.shen-traceroute-ping-ext]
Shen, N., Pignataro, C., Asati, R., Chen, E., and A. Atlas, "Traceroute and Ping Message Extension", draft-shen-traceroute-ping-ext-04 (work in progress), February 2012.

[I-D.wing-nat-reveal-option]
Yourtchenko, A. and D. Wing, "Revealing hosts sharing an IP address using TCP option", draft-wing-nat-reveal-option-02 (work in progress), June 2011.

[I-D.ymbk-aplusp]
Bush, R., "The A+P Approach to the IPv4 Address Shortage", draft-ymbk-aplusp-10 (work in progress), May 2011.

[RFC4987] Eddy, W., "TCP SYN Flooding Attacks and Common Mitigations", RFC 4987, August 2007.

Appendix A. Change History

[Note to RFC Editor: Please remove this section prior to publication.]

Author's Address

Andrew Yourtchenko
Cisco Systems, Inc.
6a de Kleetlaan
Diegem 1831
BE

Phone: +32 2 704 5494
Email: ayourtch@cisco.com